DOCUMENT RESUME

ED 170 707 CS 004 806

AUTHOR Byrne, Carolyn D.
TITLE Setting Standards and Expectations for Reading Comprehension.
PUB DATE Apr 79
NOTE 17p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, California, April 8-12, 1979); Figure 1 may not reproduce well due to small type

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Cloze Procedure; *Evaluation Criteria; Evaluation Methods; *Graduation Requirements; Minimum Competency Testing; Program Development; Readability; *Reading Ability; *Reading Comprehension; Reading Level; Reading Skills; Secondary Education; State Boards of Education; *State Standards; *Statewide Planning
IDENTIFIERS New York

ABSTRACT

In 1978 the New York State Board of Regents initiated action to establish statewide high school graduation standards for reading comprehension. The process began by measuring the readability levels of materials used in core courses and materials that adults normally read (newspapers, news magazines, and feature magazines). To do this, representative passages selected from all the types of reading materials were scaled according to the Bormuth mean cloze formula. The Board then determined that the "Degrees of Reading Power" (DRP) test series would be the official instrument for measuring student attainment, because passages in the DRP series were also scaled with the Bormuth mean cloze formula. From a comparison of all the readability levels (course materials, adult reading materials, and DRP tests), the Board set a tentative standard, requiring that all graduating students be able to read materials with a Bormuth value of .3812 at a 65% probability of success. This reading level approximates both a DRP reading passage of intermediate difficulty and the mean readability levels of 10th/11th grade course materials. Following public reaction to the proposed criteria, the Board adopted the tentative standard as the final standard in March 1979, agreeing to update the standards as additional information became available. (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED170707

Setting Standards and Expectations For Reading Comprehension

Carolyn D. Byrne
New York State Education Department
Albany, New York 12234

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Carolyn D. Byrne

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at an AERA/NCME Symposium "A Criterion-Referenced Approach to the Measurement of Reading Comprehension," San Francisco, California, April 10, 1979.

This paper is not an official document of the New York State Education Department. It simply represents the author's views on the procedure followed by the New York State Board of Regents in setting standards for reading comprehension.

CS 004 804

Setting Standards and Expectations for Reading Comprehension

Introduction

During the past several years, more than half of the states, by legislative or administrative action, have moved toward the adoption of statewide student performance standards in their elementary and secondary schools. The major impetus for the development appears to be concern about a low level of student performance in the basic skills, which typically include reading comprehension, mathematical computation, and written and oral expression. The primary sources of the concern have been those institutions that have traditionally relied upon the secondary schools to provide graduates capable of performing at adequate levels in the basic skills: colleges, universities, and the business community. Many parents, too, have expressed their concern about the apparent failure of the schools to ensure that students attain reasonable levels of proficiency in the basic skills.

The setting of standards requires that decisions be made concerning the level(s) of proficiency students need to acquire. In New York State such decisions are the prerogative of the Board of Regents. A number of states require students to demonstrate competency in the basic skills in order to receive a high school diploma. New York began moving in this direction in 1973. Initial standards were criticized as too low. Therefore, in July and August 1978, the Board of Regents introduced new tests and adopted tentative competency standards in reading comprehension, writing and mathematics.

When establishing levels of proficiency in reading comprehension, several kinds of performance standards can be defined: (1) adult standards, defined in terms of the terminal level(s) of textual difficulty that students need to be able to comprehend, (2) rational grade level standards, based on knowledge of the shape of the developmental learning curves for reaching terminal standards, (3) de facto grade level standards, defined in terms of materials assigned to students. In the absence of information on adult standards or rational grade level standards, the expectations for reading comprehension, as defined by the Board of Regents, were based on de facto grade level standards.

The materials assigned to students to be mastered during the school years constitute a set of de facto standards for reading achievement. The Board of Regents, therefore, requested a display of the readability of content area materials in use in grades 9 through 12, as well as the readability of commonly read newspapers and magazines. On the basis of this quantitative display of reading demands, the Board selected a tentative standard, expressed in terms of the readability of prose that those to be awarded a high school diploma should be able to comprehend.

The written materials that students and graduates of the educational system are expected to read were specified. These written materials were scaled for readability using the Bormuth mean cloze formula, (Bormuth, 1969). The Bormuth values were aggregated and the level of reading comprehension ability that graduates of the educational system must attain were operationally defined by the "average" measured readability characteristics of the designated materials.

Once these various quantitative standards of reading competence were defined empirically, it became possible to obtain simple and direct measurements of the extent to which students attain the standards by administering the Degrees of Reading Power Tests. The Degrees of Reading Power (DRP) Tests were designed to determine whether individuals can read with comprehension passages that have been scaled (Bormuth mean cloze formula[1]) for the same linguistic properties that were used to define the standards for readability.

The Degrees of Reading Power

The tests called the Degrees of Reading Power are designed to be a sensitive measure of the ability to read with comprehension.[2] They consist of a series of non-fiction prose passages on randomly selected topics. The passages are presented in order of difficulty beginning with very easy text and progressing by degrees to very difficult text. Test items are created by the deletion of seven

---

[1] "This formula, selected following a critical review of the literature, was chosen for its technical superiority in several respects: (a) the Bormuth formula yields a set of measurements with at least interval properties, a requirement for an effectiveness measure: (b) Bormuth used cloze tests to derive criterion scores in developing this formula, and the literature strongly suggests that cloze scores, a measure of redundancy in language, are the most objective scores to use for this purpose, in that cloze deletions ensure prima facie that the difficulty of questions is tied to the difficulty of the material; (c) it provides a continuous scale over the entire range of readability of first through at least twelfth grade instructional materials; (d) it has a higher validity coefficient ($\bar{r}$=.83 in original development and $\bar{r}$=.92 in cross-validation) than other readability formulas; (e) it has relatively low standard errors of measurement (S_{em}=.06 in original development and S_{em}=.04 in cross-validation)." (Koslin, Koslin, and Zeno, 1978a, p.36)

[2] "The DRP defines reading comprehension as the ability to use the syntactic and semantic information in a mutilated prose text to complete the message correctly. The DRP measures the ability to derive sufficient meaning(s) from connected prose text to correctly complete the message in adjacent text that has been mutilated. The test measures not only understanding of the messages within individual sentences, but also the ability to integrate meaning across sentences and to understand the more extended messages denoted in the text. The DRP is designed to measure the ability of an individual to derive such meanings from any non-technical expository text at different levels of difficulty" (Koslin, et. al., 1978a, p. 28).

words in each passage, with the student selecting the most appropriate word from the five options provided for each deletion.

The DRP can be criterion-referenced; scores on the measure can be interpreted with respect to a criterion, provided that the criterion is expressed in terms of the readability of prose text. Thus, if the materials that students need to be able to read are identified and analyzed, the readability of these materials constitutes the standard or criterion whose attainment the DRP can evaluate. (Koslin, et. al., 1978a)

"The purposes of the DRP require that item difficulty be directly linked to passage difficulty. Standard item-writing techniques do not ensure that test items accurately reflect the difficulty of textual material. However, cloze procedures do generate items which are prima facie tied to the difficulty of the text in which they appear. Because DRP items are somewhat related psycholinguistically to cloze, and because of the rules of formation for the items, DRP items reflect the difficulty of the text in which they appear" (Koslin, 1978, p. 37).

Since "average" Rasch item difficulty per passage highly correlates with linguistic features of the text known to account for its difficulty and since "average" Rasch item difficulty per passage highly correlates with text readability, the DRP item difficulties have empirical meaning with respect to prose that students can understand. One regression equation allows DRP ability scores to be converted into forecasts of the prose that students can read. The other regression equation allows for the prediction of the ability required to comprehend prose of a given level of readability.

The regression equation predicting DRP difficulties from readability enables the DRP to be "objectively" criterion-referenced. Thus, if readability formulas are used to analyze a set of reading tasks, it is possible to use the results of these analyses to find the DRP passage difficulties which correspond in readability to minimal, usual (average), and any other adult prose reading tasks. (Koslin, 1978a)

The correlation between the ordinal position of a passage on the DRP and its mean and median passage difficulty is essentially one. Since the DRP is constituted by writing passages at a pre-assigned level of readability using the Bormuth mean cloze formula, the correlation between Bormuth cloze scores and ordinal position must be one. (Koslin, 1978a)

- -

Insert Table 1 about here

- -

Sampling to Define De Facto Reading Standards

The materials assigned to students to be mastered during the school years constitute the set of expectations or de facto reading standards. In order to define these de facto reading standards in grades 9-12, a statement concerning textbook usage was required. Carroll, Davies, and Richman (1971) selected a sample to define (national) de facto reading standards. The corpora compiled by Carroll, et. al. (1971) were not suitable because: (a) the list does not contain information about textbook usage in grades 10 to 12, and (b) no evidence is presented to warrant the conclusion that the list for grade 9 represents materials used in New York State.

The plan used by the New York State Education Department for selecting materials for the readability analysis is discussed under three separate categories: textbooks, magazines and newspapers.

Textbooks

In defining grade level corpora of instructional materials in grades 9-12, it is unnecessary to analyze the readability of every book used in the schools. Books that are used for supplementary study and resource(s) purposes in the classroom were excluded since many of them are used by few rather than all students and thus do not reflect stable reading demands that schools are placing on students. Library books were also excluded, since it is impossible to estimate accurately how many students use them or the grade(s) in which they are used.

The only textbooks on which an analysis was necessary were those used in the core subjects. The core subjects are those subjects in which all students in New York State Schools must obtain a passing grade in order to receive a high school diploma: English (4 units), social studies (3 units), science (1 unit) and health (½ unit).

The subject matter specialists at the New York State Education Department were sources of information of the textbooks commonly in use in the secondary schools. Thus, these professionals were asked to compile a list of the most commonly used textbooks in the subjects for which they had primary responsibility.

To determine the readability of each appropriate[3] textbook listed, three separate data points of approximately 250 words in length were randomly selected from within the fifth tenth of the book. The Bormuth formula was applied to these data points, with the results as shown.

---

[3]A textbook was designated as appropriate if it consistently contained 250 words of continuous prose text.

- -

Insert Figure 1 about here

- -

Magazines

The list of magazines to be sampled was taken from The World Almanac and Book of Facts 1975. The Almanac used as its source the Audit Bureau of Circulations FAS-FAX Report. Listed were general and farm magazines, exclusive of groups and comics, based on total average paid circulation during the six months prior to December 31, 1972.

This list, containing 101 magazines, was divided into five clusters. The clusters were based on circulation, as shown:

Cluster Number	Circulation
1	more than 3 million but less than 19 million
2	more than 1.5 million but less than 3 million
3	more than .75 million but less than 1.5 million
4	more than .50 million but less than .75 million
5	more than .40 million but less than .50 million

From the list of 101 magazines arranged in five clusters according to circulation, every other magazine was selected. If a particular magazine was not available, the next available magazine listed was selected (Weight Watchers was substituted for American Girl, Mechanics Illustrated for Elks Magazine, Motor Trend for Flower and Garden, and Skiing for Cappers Weekly). The magazines in each cluster were divided into newsmagazines and feature magazines. A magazine was defined as a newsmagazine if its primary content was current events, economic and/or business news; all other magazines were classified as feature.

The articles within the news magazines, by clusters, were divided into five categories: editorials, sports, international news, national news, regional news. After the articles were categorized, each article containing a minimum of 250 words was assigned a number by category. Three numbers were then drawn from a hat for each category. The articles assigned these numbers constituted the sample upon which the readability analysis of news magazines was based. The results are as shown in Figure 1.

In each cluster of feature magazines, each article of at least 250 words was assigned a number. Three numbers were then randomly selected within each cluster. All magazines used were the most current, on the newstands as of June 5, 1978. The results are shown in Figure 1.

Newspapers

The list of newspapers to be sampled was taken from The World Almanac and Book of Facts 1975; the list entitled "Selected United States Daily Newspapers' Circulation" was used. The source for this list was the Audit Bureau of Circulations FAS-FAX Report of average paid circulation for the six months ending March 31, 1974. Sunday newspapers were omitted from the sample and only New York State newspapers were used.

These New York State newspapers were divided into five clusters, based on circulation as shown:

Cluster Number	Circulation
1	more than 1 million
2	more than .50 million but less than 1 million
3	more than .25 million but less than .50 million
4	more than .10 million but less than .25 million
5	more than .06 million but less than .10 million

The articles within the newspapers, by clusters, were divided into five categories: editorials, sports, international news, national news, regional news. After the articles were categorized, each article containing a minimum of 250 words was assigned a number by category. Three numbers were then drawn from a hat for each category. The articles assigned these numbers constituted the sample upon which the readability analysis of newspapers was based.

All newspapers used were the most current on the newstands as of June 2, 1978, with the exception of the Rochester Democrat Chronicle.

Figure 1 is a representation of the reading difficulty of passages which appear on a Degrees of Reading Power Test and the reading difficulty of selected passages from the textbooks, newspapers and magazines sampled. The DRP passages are listed in order of difficulty beginning with the easiest (Deserts with a Bormuth value of .5726) and progressing to the most difficult (Synchrotron Emissions with a Bormuth value of .1690).

The mean difficulties of the articles from the news and feature magazines were designated by (▲), while the mean difficulties of New York newspapers were indicated by the (○). It can be seen that the readability level of these sources is relatively homogeneous.

As is evident in Figure 1 the mean readability values for the textbooks are lower (more difficult for each succeeding grade, 9-12).[4] The simplest mean readability value in the sample(s) corresponds approximately to the seventh passage in the DRP, .4023. The hardest, or 12th grade materials in the sample have a mean readability about equal to the tenth passage on the DRP, .3388.

[4]Expressed in terms of a widely used readability formula (Dale-Chall), the mean values in the sample correspond roughly to grades 5 to 17.

The analysis of the readability of books in modal use in the core subjects provides objective information concerning what students are expected to be able to read in grades 9-12. Each new grade begins with text whose mean readability is somewhat harder than the text was at the beginning of the previous grade.

The analysis shows that the de facto standard in reading for students in the ninth grade has a Bormuth cloze mean readability value of .4116. The analysis also shows that de facto standards in reading for students in the twelfth grade has a Bormuth cloze mean readability value of .3405.

The Board of Regents, using in part the information from the readability analysis of the textbooks in modal use, set the average readability of 10th and 11th grade subject area materials as a tentative standard for high school graduation.

The tentative standard which the Regents selected is equal in difficulty, at a 65% comprehension level, to the DRP passage on Anesthetics (Bormuth readability of .3812). The DRP was adopted as the method of measuring whether the standard is attained. "In every DRP test used to demonstrate achievement each student shall be required to attain a grade of 65 percent, that is to demonstrate that (s)he can read 10th grade and 11th grade material with a 65 percent probability of success." (Board of Regents Resolution, July, 1978.)

Given the crucial importance of this decision the Board actively sought large scale public reaction to the proposed standard. The representative community-based participation the decision-making process was gained through regional conferences, newspaper surveys and media coverage. In addition in December, 1978, copies of the Information Brochure for the Regents Competency Testing Program were sent to a 20 percent random sample of schools. Each member of the staff was asked to complete a questionnaire indicating the minimum standards that should be required for a high school diploma. The sample included both elementary and secondary schools, public and nonpublic. (See Appendix A for questionnaire.)

- -

Insert Table 2 about here

- -

Based on this information, the Board of Regents, at its March 1979 meeting, adopted the tentative standard as the final standard. This standard defines a floor or minimum level of reading comprehension for any student being processed by the educational system. The New York State Board of Regents set a standard requiring that all graduating sutdents be able to read

materials with a Bormuth value of .3812 at a 65 percent probability of success. Therefore, the schools must turn out students who can read material at that level of difficulty.

Forecasting Future Reading Requirements

For present purposes the New York State Education Department has not determined the readability of any systematic or representative sample of reading tasks that adults actually engage in or ought to be able to achieve. Such work lies in the future as part of the setting of rational standards.

The setting of rational standards will be based upon the reading requirements that citizens will need to meet fifteen to twenty years in the future. The program of reading instruction that a student receives from the time he enters school should be geared to preparing him for standards he will need to meet, at the time of high school graduation, in order to enter college, technical school, or an apprenticeship program to get a job, or to function effectively in the adult world in general.

To this end the Department is presently committed to performing readability analyses on adult materials. This commitment is based on the fact that adults must read more than they had to in earlier times (Fries, 1962). Each year there are fewer jobs that can be filled by persons with limited reading skills. These changes have resulted in a sharp increase in the reading demands on adults over the last generation. Therefore, to do an adequate job of educating students to be competent adult readers, it is necessary to be able to forecast the reading skills that a student entering first grade now will need by the time he is a graduating twelfth grader.

As additional information becomes available, the setting of standards for reading competency can be updated and based on informed decision.

Table 1

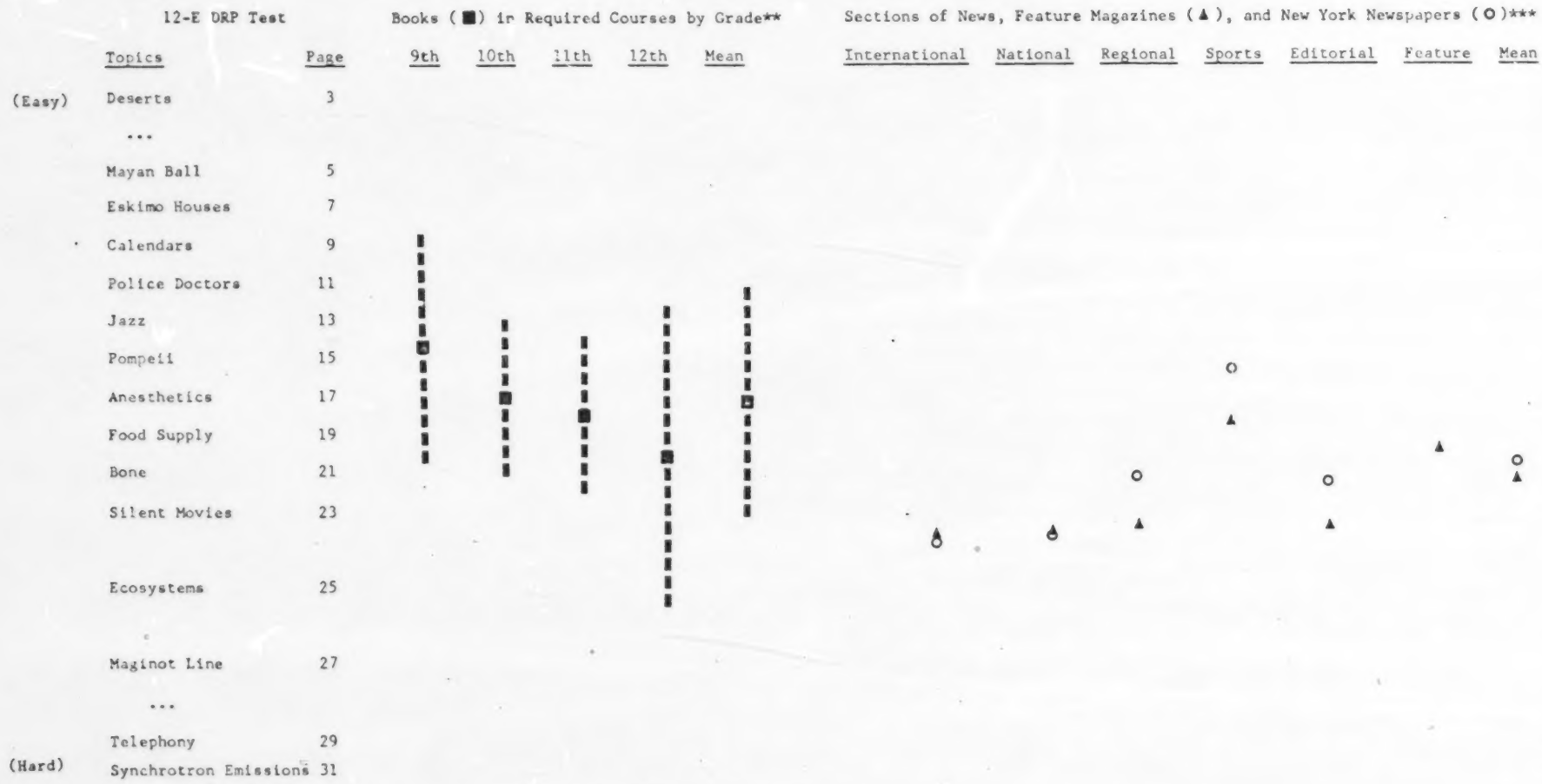
Intercorrelations Between Mean and Median DRP Passage Difficulties and Bormuth Readability Values[a]

	Bormuth	Ordinal Position DRP Passage
Form Y		
Mean DRP Passage Difficulties	.980	.980
Median DRP Passage Difficulties	.971	.961
Bormuth Cloze Scores		1.000
Form X		
Mean DRP Passage Difficulties	.944	.944
Median DRP Passage Difficulties	.965	.965
Bormuth Cloze Scores		1.000

Note: [a]Mean cloze scores

Note: From "The Validity and Reliability of the Degrees of Reading Power Test" by B.L. Koslin, S. Koslin, H. Wainer, and S. Zeno. Touchstone Applied Science Associates, 1978. Reprinted by permission.

Figure 1
Text Difficulty: Means and Spread*



Notes: *Based on 250 word samples
**96 Samples weighted according to requirements: English and Social Studies texts sampled for all grades; (Social Studies included in 12th grade sample since it is the dominant elective); Science texts in grade 9, and Health texts in grade 11.
***56 magazine and 75 newspaper samples were analyzed by circulation. Neither the circulations of the News nor Feature Magazines nor Newspapers made a difference.

Table 2

Median Scale Values of Recommended Standards for Selected Groups

Group	Subject (Reading Comprehension)
Board of Regents	5.0
Newspaper Respondents	5.8
Regional Conference Participants	4.8
Secondary English Teachers	5.5
Secondary Math Teachers	5.4
All Secondary Teachers	5.6
Professional Organizations and others	5.5

Note: From "High School Diploma Standards Recommended by School Personnel" by W.A. Lott, The University of the State of New York, The State Education Department, 1979. Reprinted by permission.

Appendix A

QUESTIONNAIRE

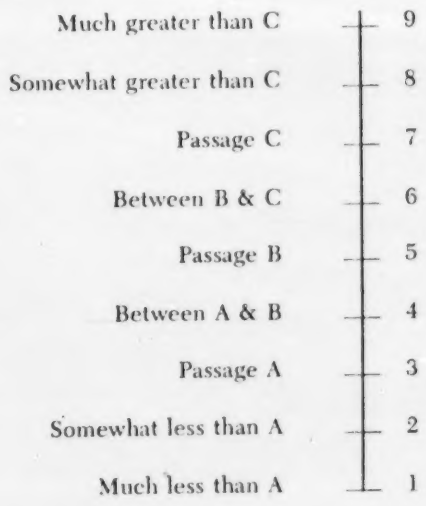
Please indicate your views on what the standards should be in each skill area for graduation from high school in New York State by circling the appropriate number on each of the scales below.

Reading Comprehension

What is the minimum difficulty level of materials that students receiving a high school diploma should be able to read with comprehension?

Please circle the number at the point on the scale at the right that best corresponds to your view. For example, if you feel that Passage A represents an appropriate minimum, circle the number 3; if you think that an appropriate minimum would be between Passages B & C, circle the number 6; if you believe that an appropriate minimum would be somewhat less than Passage A, circle the number 2; etc.

Reading Comprehension


Much greater than C 9
Somewhat greater than C 8
Passage C 7
Between B & C 6
Passage B 5
Between A & B 4
Passage A 3
Somewhat less than A 2
Much less than A 1

References

Bormuth, F.R. "Readability: A New Approach." Reading Research Quarterly, 1966, 1, 79-132.

Bormuth, J.R. "Cloze Test Readability: Criterion Reference Scores." Journal of Educational Measurement. 1968, 5, 189-196.

Bormuth, J.R. Development of Readability Analyses. Final report, Project No. 7-0052, Contract No. OEC-3-7-070052-0326. Office of Education, Bureau of Research, U.S. Department of Health, Education, and Welfare. March, 1969. (a)

Bormuth, J.R. "Factor Validity of Cloze Tests as Measures of Reading Comprehension Ability." Reading Research Quarterly. 1969, 4, 358-365. (b)

Bormuth, J.R. Development of Standards of Readability: Toward a Rational Criterion of Passage Performance. Final Report, Project No. 9-0237, Grant No. OEG-0-9-230237-4125, Bureau of Research, Office of Education, U.S. Department of Health, Education, and Welfare, 1971.

Carroll, J.B., Davies, P., & Richman, B. The American Heritage Word Frequency Book. Boston and New York; Houghton Mifflin Co. and American Heritage Publishing Co., 1971.

Dale, E. & Chall, J. "A Formula for Predicting Readability." Educational Research Bulletin. 1948, 27(1), 11-20, 28; 27(2), 11-28.

Fries, C.C. Linguistics and Reading. New York: Holt, Rinehart, and Winston. 1962.

Koslin, S., Koslin, B.L., Langley, C., & Josephson, B. Design Concepts for A Measure of Effectiveness in Reading. New York, Riverside Research Institute, 1973.

Koslin, S., Koslin, B.L. "The Further Development of an Effectiveness Measure in Reading." New York, Touchstone Applied Science Associates, 1978. (a)

Koslin, S., Koslin, B.L., Wainer, H., Zeno, S. "The Validity and Reliability of the Degrees of Reading Power Test." Touchstone Applied Science Associates, 1978. (b)

Kyte, G.C. "Calibrating Reading Material." Elementary School Journal, 1925, 25, 533-546.

Lott, W.A., "High School Diploma Standards Recommended by School Personnel." New York State Education Department, 1979.

"The Regents Competency Testing Program Information Brochure." New York State Education Department, 1979.